On almost sure identifiability of non multilinear tensor decomposition

Jérémy E Cohen, Pierre Comon

▶ **To cite this version:**

Jérémy E Cohen, Pierre Comon. On almost sure identifiability of non multilinear tensor decomposition. EUSIPCO 2014 - 22th European Signal Processing Conference, Sep 2014, Lisbonne, Spain. pp.2245-2249. hal-01059060v2

HAL Id: hal-01059060
https://hal.science/hal-01059060v2

Submitted on 17 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON ALMOST SURE IDENTIFIABILITY OF NON MULTILINEAR TENSOR DECOMPOSITION

Jeremy E.Cohen and Pierre Comon

Gipsa-Lab, CNRS
Grenoble Campus, BP.46,
F-38402 St Martin d'Heres cedex, France

ABSTRACT

Uniqueness of tensor decompositions is of crucial importance in numerous engineering applications. Extensive work in algebraic geometry has given various bounds involving tensor rank and dimensions to ensure generic identifiability. However, most of this work is hardly accessible to non-specialists, and does not apply to non-multilinear models. In this paper, we present another approach, using the Jacobian of the model. The latter sheds a new light on bounds and exceptions previously obtained. Finally, the method proposed is applied to a non-multilinear decomposition used in fluorescence spectrometry, which permits to state generic local identifiability.

1. INTRODUCTION

The usefulness of tensors, which are represented by multiway arrays, has been growingly recognized in the last decades, as testified by some surveys including [1, 2]. In particular, chemometrics make extensive use of tensor decompositions. For example, in order to detect dangerous chemical components in water, one can use fluorescence spectrometry, measuring the components' concentration by building Fluorescent Excitation-Emission Matrices (FEEM) of the solution [3]. More precisely, if inner effects are neglected, and if Raman and Rayleigh diffusion phenomena are removed, the intensity recorded in the stacked FEEM matrices can be assumed to follow rather accurately the model below:

$$\boldsymbol{\mathcal{G}}(\boldsymbol{\theta}) = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r \tag{1}$$

where vectors $\mathbf{b}_r$ and $\mathbf{c}_r$ represent respectively the excitation and emission spectra of the rth solute, $\mathbf{a}_r$ contains the relative concentration of the rth solute in every experiment, and $\boldsymbol{\theta}$ is a parameter vector containing all vectors $\{\mathbf{c}_r, \mathbf{b}_r, \mathbf{a}_r\}$. All these quantities are a priori unknown. In other words, we measure tensor coordinates stored in an array $\boldsymbol{\mathcal{T}}$, which follows a model $\boldsymbol{\mathcal{T}} \approx \boldsymbol{\mathcal{G}}(\boldsymbol{\theta})$, and the goal is to identify $\boldsymbol{\theta}$ from the sole observation of $\boldsymbol{\mathcal{T}}$. This problem is not specific to chemometrics, and can be encountered in a wide panel of applications [1].

This work has been funded by the European Research Council under the European Communitys Seventh Framework Programme FP7/20072013 Grant Agreement no. 320594.

When the integer R in (1) is minimal, it is called the rank of tensor $\boldsymbol{\mathcal{G}}$ and is denoted $rank_{\otimes}(\boldsymbol{\mathcal{G}})$. We are actually interested in a low-rank approximation of $\boldsymbol{\mathcal{T}}$, as in most applications the latent information is compact and polluted by noise, which increases the exact rank.

Mathematically speaking, the decomposition (1), often referred to as CP decomposition, is non trivial to compute [4, 2] for the following reasons. There exist at least two differences between matrices and tensors of order strictly larger than 2 [4, 2]. Firstly, a best low-rank tensor approximate does not always exist. Secondly, the exact CP tensor decomposition may be unique under sufficient conditions (contrary to matrices, for which there exist an infinity of such decompositions).

Kruskal was the first to address the issue in his renowned publication [6], and gave an upper bound on tensor rank to ensure global uniqueness. Tighter bounds can be found in [7, 9], which cover the n-way case as well. On the other hand, the problem of existence of a best approximate of rank $r < rank_{\otimes}(\boldsymbol{\mathcal{T}})$ is well summarized in [4, 8].

In this paper, we address the question of generic local uniqueness of decomposition (1) or its approximation, in other words, local identifiability of a tensor CP model. In algebraic geometry, computing (1) is known as the Waring problem, and has been addressed in particular by the italian school [10, 11]. Surprising results are found for R equal or very close to what is called the expected rank of $\boldsymbol{\mathcal{T}}$, that is, the rank that we expect $\boldsymbol{\mathcal{T}}$ to have, considering its dimensions and the fact that it is drawn randomly according to an absolutely continuous distribution. For engineers, the most interesting result is that if tensor rank is strictly smaller than the expected rank minus one, the CP identifiability is generically verified [11, 5].

Despite its power, the algebraic geometry approach of identifiability has drawbacks. First, it does not apply to non multilinear models, as described in Section 4. Moreover, a strong mathematical background is necessary to understand these contributions. Hence we suggest a simple constructive approach for studying identifiability, whose results are globally not new, but which extends beyond the scope of linear spaces, and offers an easy understanding. The latter has been also used in [12] to compute numerical simulations, but no theoretical analysis was made therein.

In the next section we consider the canonical decomposition model given in (1), and describe the Jacobian of the model

for rank 1, and then for any rank. Section 3 exposes our constructive approach for detecting exceptions to the general identifiability rule of model (1). Section 4 contains our main result, which extends previous ones to non multilinear decompositions, as that described in [3].

2. IDENTIFIABILITY, A CONSTRUCTIVE APPROACH

As explained in the introduction, the CP model consists of a sum of rank one tensors. Identifiability means uniqueness of these rank one terms. A standard way to check local uniqueness of a model is to determine whether the Jacobian of the model is at least partially invertible (we will see in what sense later) or not. In the following, we limit ourselves to 3-way arrays, but the discussion can be easily extended to d-way arrays.

Let us store the entries of an array $T_{ijk} \in \mathbb{R}^{K\times L\times M}$ in a vector of coordinates, $\mathbf{x}$, of size KLM. Also note $\mathbf{a}_r$, $\mathbf{b}_r$ and $\mathbf{c}_r$, the columns of respectively $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, and concatenate them at will in a single vector $\boldsymbol{\theta}$ of size $R(K+L+M)$. The goal is to fit a model $\boldsymbol{\mathcal{G}}(\boldsymbol{\theta})$ to the observed vector $\mathbf{x}$. *Local identifiability* is related to the invertibility of function $\boldsymbol{\mathcal{G}}(\cdot)$ in an open neighborhood of $\boldsymbol{\theta}$ by the local inversion theorem. Since function $\boldsymbol{\mathcal{G}}(\boldsymbol{\theta})$ is differentiable, we have the first order relationship:

$$d\mathbf{x} = \mathbf{J}(\boldsymbol{\theta})\, d\boldsymbol{\theta}$$

where $\mathbf{J}(\boldsymbol{\theta})$ denotes the Jacobian of $\boldsymbol{\mathcal{G}}(\boldsymbol{\theta})$, *i.e.* $J_{ij} = \partial \boldsymbol{\mathcal{G}}_i/\partial\theta_j$. If $\mathbf{J}$ is full column rank, then we can compute the modified parameter vector $\boldsymbol{\theta} + d\boldsymbol{\theta}$ from the modified observed data $\mathbf{x} + d\mathbf{x}$. In particular, covariance matrices are related to each other as $\mathbf{V}_x = \mathbf{J}\,\mathbf{V}_\theta\,\mathbf{J}^T$, which yields $\mathbf{V}_\theta = \mathbf{J}^\dagger\,\mathbf{V}_x\,\mathbf{J}^{\dagger T}$, where $(^\dagger)$ denotes the Moore-Penrose pseudo inverse. We can observe that since $\mathbf{V}_x$ is given, the conditioning of $\mathbf{V}_\theta$ is controlled by $\mathbf{J}$.

At this stage, a link can be established with the Maximum Likelihood (ML) approach. If observation $\boldsymbol{\mathcal{T}}$ follows model (1) up to an additive Gaussian noise of covariance matrix $\mathbf{W}$, *i.e.*:

$$\mathbf{x} = \boldsymbol{\mathcal{G}}(\boldsymbol{\theta}) + \boldsymbol{\epsilon} \tag{2}$$

then the Fisher information matrix is given by $\mathbf{F}(\boldsymbol{\theta}) = \mathrm{E}\{\mathbf{s}(\mathbf{x},\boldsymbol{\theta})\,\mathbf{s}(\mathbf{x},\boldsymbol{\theta})^T\}$, where $\mathbf{s}(\mathbf{x},\boldsymbol{\theta}) = \partial||\mathbf{x} - \boldsymbol{\mathcal{G}}(\boldsymbol{\theta})||^2_W/\partial\boldsymbol{\theta}$ is the score function. It can be easily shown that this leads to $\mathbf{F}(\boldsymbol{\theta}) = \mathbf{J}^T\,\mathbf{W}^{-1}\mathbf{J}$. In other words, when $\mathbf{W}$ is invertible, the Cramér-Rao (CR) bound is finite if and only if the Jacobian $\mathbf{J}$ is full column rank. When $\mathbf{W} = \sigma^2\mathbf{I}$, CR bounds for the CP decomposition problem can be found in [13, 14].

This *local identifiability* based on the conditioning of matrix $\mathbf{J}(\boldsymbol{\theta})$ can be checked either at a given point $\mathbf{x}$, or at random points. If local identifiability is satisfied almost everywhere, for random points drawn according to an absolutely continuous probability distribution, then we have *generic local identifiability* of the model. For the CP model, we set:

$$\mathbf{g}(\boldsymbol{\theta}) = \sum_{r=1}^{R} \mathbf{a}_r \boxtimes \mathbf{b}_r \boxtimes \mathbf{c}_r \text{ and } \boldsymbol{\theta} \stackrel{\text{def}}{=} vec\{\begin{bmatrix}\mathbf{C}\\ \mathbf{B}\\ \mathbf{A}\end{bmatrix}\} \stackrel{\text{def}}{=} \begin{pmatrix}\mathbf{c}_1\\ \mathbf{b}_1\\ \mathbf{a}_1\\ \vdots\\ \mathbf{a}_R\end{pmatrix}$$

where $\boxtimes$ denotes the Kronecker product. Differentiating $\mathbf{g}$ yields the expression of the $KLM \times (K+L+M)R$ Jacobian matrix $\mathbf{J}$, which can also be found in [12, 15]:

$$\mathbf{J} = [\mathbf{a}_1 \boxtimes \mathbf{b}_1 \boxtimes Id_M \mid \mathbf{a}_1 \boxtimes Id_L \boxtimes \mathbf{c}_1 \mid \ \ldots\ \mid Id_K \boxtimes \mathbf{b}_R \boxtimes \mathbf{c}_R]$$

As an illustration, we give below a schematic description of $\mathbf{J}$, *i.e.* with $(L, M, K) = (2, 2, 2)$ and $R = 1$.

$$\mathbf{J} = \begin{pmatrix} a_1b_1 & 0 & a_1c_1 & 0 & b_1c_1 & 0\\ 0 & a_1b_1 & a_1c_2 & 0 & b_1c_2 & 0\\ a_1b_2 & 0 & 0 & a_1c_1 & b_2c_1 & 0\\ 0 & a_1b_2 & 0 & a_1c_2 & b_2c_2 & 0\\ a_2b_1 & 0 & a_2c_1 & 0 & 0 & b_1c_1\\ 0 & a_2b_1 & a_2c_2 & 0 & 0 & b_1c_2\\ a_2b_2 & 0 & 0 & a_2c_1 & 0 & b_2c_1\\ 0 & a_2b_2 & 0 & a_2c_2 & 0 & b_2c_2 \end{pmatrix}$$

Now this matrix contains all the information we need about local uniqueness, and thus we already know that finding its rank will be non-trivial, because of surprising exceptions that have been already pointed out (see Section 3). Yet, we are able to dig further in the matter with a few propositions.

The next proposition is a simple matrix interpretation of the scaling indeterminacies in Equation (1). A necessary condition for a matrix to admit a left inverse is that it has at least as many rows as columns. And for a random matrix drawn according to an absolutely continuous distribution, this condition becomes also sufficient. We shall say that this condition is *generic*. For the Jacobian matrix, from the scaling indeterminacies stem two linear relations between the columns within every $K + L + M$ block, so that we have the expected property in terms of Jacobian rank:

Proposition 1 *For a generic tensor $\boldsymbol{\mathcal{T}}$ in $\mathbb{C}^{K\times L\times M}$,*

$$rank(\mathbf{J}) \leq (K + L + M - 2)R$$

This result is well known. In fact, loading matrices are generically full rank, but scaling indeterminacies remain in expressions of rank-1 terms, involving $2R$ free parameters [6].

We call $\overline{R}$ the critical value of tensor rank defined by

$$\overline{R} = \left\lfloor \frac{KLM}{K + L + M - 2} \right\rfloor \tag{3}$$

Note that $\overline{R}$ is defined by flooring the ratio, whereas the *expected rank* is usually defined by ceiling it.

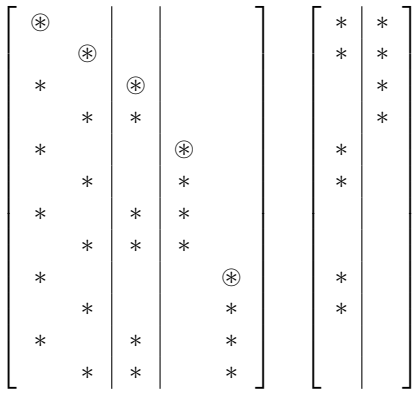
Proof. Denote by $\mathcal{C}_i$ the i^{th} column of matrix $\mathbf{J}$ previously defined. The following two relations are verified with indices

modulo $K+L+M$:

$$\sum_{i=M+1}^{M+L} \mathcal{C}_i = \sum_{j=1}^{M} \mathcal{C}_j, \qquad \sum_{i=M+L+1}^{M+L+K} \mathcal{C}_i = \sum_{j=1}^{M} \mathcal{C}_j$$

Hence there are at most $R(K+L+M-2)$ linearly independent columns. □

Another result on CP identifiability is that a rank-1 approximation is always locally unique. In fact, this can be seen on the Jacobian matrix, as the previous proof reveals a staircase structure in the matrix for rank 1 approximation, if one cancels out the $(M+1)^{th}$ and $(L+M+1)^{th}$ columns in every $K+L+M$ block. As an example, the case where $(L, M, K) = (2, 2, 3)$ is depicted below: two columns are isolated and deleted (right) yielding a triangular matrix (left).



If the circled coefficients in the figure above are not null, we have then extracted a triangular invertible square submatrix of size $K+L+M-2$ from the Jacobian $\mathbf{J}$. This will be the case if tensor $\boldsymbol{\mathcal{T}}$ is generic. This shows that under this condition, there is locally a unique solution; in other words, we have proven local identifiability, which leads to the following proposition.

Proposition 2 *For $R = 1$ and any tensor $\boldsymbol{\mathcal{T}}$ satisfying the noiseless model (1) generic local identifiability is ensured.*

This proposition holds true for the noisy model (2) in the presence of small noise. In addition, from Bézout Theorem, the number of solutions minimizing the Frobenius norm $||\mathbf{x}-\boldsymbol{\mathcal{G}}(\boldsymbol{\theta})||^2_W$ is generically finite (and bounded by $5^{K+L+M-2}$ when $R = 1$), since stationary points are defined by a system of polynomial equations. Unfortunately, Proposition 2 does not extend to ranks higher than 1, as discussed in the next section.

3. EXCEPTIONS TO IDENTIFIABILITY

Up to now, we have exposed generic results on tensor CP decomposition, which can be intuitively interpreted with the Jacobian conditioning. We defined the critical rank (3) of the exact decomposition of a generic tensor as the largest integer R that allows the Jacobian of the multilinear model (1) to have more lines than columns. That is, the critical rank maximizes the number of unknowns while leaving more equations than unknowns, and we thus expect such a model to be identifiable. In other words, we expect equality to occur generically for all K, L, M in the inequality of Proposition 1, and the Jacobian to be full rank up to the scaling indeterminacies for a decomposition of rank R smaller than the critical rank.

However, in some exceptional settings, local identifiability is generically not achieved by the CP decomposition model. This means that linear relations between columns stem from the structure of the Jacobian matrix, and unknowns stored in $\boldsymbol{\theta}$ cannot be retrieved from the measurements $\boldsymbol{\mathcal{T}}$ even though the system has more equations than unknowns. These exceptions to identifiability of the CP decomposition have been studied using algebraic geometry tools in [9, 10], and we summarize the known results that apply in our setting in the table below, taken from [9]. Note that the so-called weak defectivity defined therein does not mean that local uniqueness is not verified, and is hence not reported below.

	Defective unbalanced
Dimensions	$(K-1)(L-1)+3 \leq M$
Rank	$(K-1)(L-1)+2 \leq R$ and $R < \min(M, KL)$

	Defective
Dimensions	$K = L = 4$ and $M = 3$
Rank	$R = 5 = \overline{R}$

Because a lack of identifiability infers a non left invertible Jacobian, we should be able to give constructive proofs of the above exceptions. However, even in the simplest defective case, finding the exact symbolic expressions of the diagonal elements of the triangular form of the Jacobian has shown to be inextricable. Nevertheless, a computer simulation easily confirms the degenerescence of the Jacobian in the defective cases reported above. In this sense, algebraic geometry proves to be a more powerful theoretical tool for studying identifiability, but the Jacobian yet holds all the needed information.

4. A NON MULTILINEAR DECOMPOSITION

All the previously exposed results did not improve existing bounds on CP identifiability, and we merely described a constructive alternative approach. Now, our claim is that our method is – to our knowledge – the most efficient way to study more exotic tensor decompositions, as subsequently demonstrated.

4.1. Model

We are interested in a decomposition called Non Linear FEEM Decomposition (NLFD), which appears in chemometrics [3]. It is given by the following equation.

$$[\boldsymbol{\mathcal{G}}_{NL}(\boldsymbol{\theta})]_{ijk} = [\boldsymbol{\mathcal{G}}(\boldsymbol{\theta})]_{ijk} \prod_{p=1}^{R} e^{-\mu\left[a_{ip}(b_{jp}+b'_{kp})\right]} \qquad (4)$$

where μ is a given constant, and vectors $\mathbf{b}'_r$ are deduced from $\mathbf{b}_r$ by truncation or zero-padding, depending on L and M: $b_{kr} = b_{jr}$ for $k \leq \min(L, M)$, and if $L < M$, then $b_{kr} = 0$ for $L < k \leq M$. In [3], the definition of $\mathbf{b}'$ offers the possibility to include a shift, which we shall not take into account for the sake of simplicity, without loss of generality. As seen in Section 2, it is well known that CP loadings $\{\mathbf{a}_r, \mathbf{b}_r, \mathbf{c}_r\}$ are identifiable up to scaling indeterminacies. This result is slightly different in the NLFD case. Denote by $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ three scaling factor vectors of size R, such that:

$$\boldsymbol{\mathcal{G}}_{NL}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{G}}_{NL}(\boldsymbol{\theta}'), \text{ with } \boldsymbol{\theta}' \stackrel{\text{def}}{=} vec\{\begin{bmatrix} \mathbf{C}Diag(\boldsymbol{\gamma}) \\ \mathbf{B}Diag(\boldsymbol{\beta}) \\ \mathbf{A}Diag(\boldsymbol{\alpha}) \end{bmatrix}\} \tag{5}$$

By differentiating (5) with respect to variables $\mathbf{a}_r$ and $\mathbf{c}_r$, one can prove that $\alpha_r\beta_r\gamma_r = 1, \ \forall r$. Then, differentiation w.r.t. $\mathbf{b}_r$ yields $\forall r, 1 \leq r \leq R$:

$$\alpha_r\beta_r = 1, \quad \gamma_r = 1$$

As a consequence, the scaling indeterminacy is removed in the third mode (emission spectra in the frame of fluorescence spectroscopy). This leads to a new critical rank $\overline{R}$ for this decomposition:

$$\overline{R} = \left\lfloor \frac{KLM}{K + L + M - 1} \right\rfloor$$

Regarding identifiability, the situation is the same as for the multilinear case (Section 2). We want to fit model $\boldsymbol{\mathcal{G}}_{NL}(\boldsymbol{\theta})$ given in equation (4) to the data vector $\mathbf{x}$. Let us first define the KLM-dimensional vectors:

$$\begin{aligned} \mathbf{e} &= \mu \sum_{r=1}^{R} \mathbf{a}_r \boxtimes (\mathbf{b}_r \boxtimes \mathbf{1}_M + \mathbf{1}_L \boxtimes \mathbf{b}_r) \\ \mathbf{h} &= \exp(-\mathbf{e}) \end{aligned}$$

where the exponential acts entry-wise. Differentiating $\boldsymbol{\mathcal{G}}_{NL}$, we find the following measurement relation between the data $\mathbf{x}$ and the model variables $\boldsymbol{\theta}$:

$$d\mathbf{x} = d\mathbf{g}_{NL}(\boldsymbol{\theta}) = \mathbf{h} \boxdot \left(\mathbf{J}\, d\boldsymbol{\theta} - \mathbf{g}(\boldsymbol{\theta}) \boxdot d\mathbf{e}(\boldsymbol{\theta})\right) \tag{6}$$

where $\mathbf{J}$ is the Jacobian of the CP decomposition defined in Section 2 and $\boxdot$ is the component-wise (so-called Hadamard) product. In (6), the differential of the NL term takes the form:

$$\begin{aligned} d\mathbf{e}(\boldsymbol{\theta}) = \mu \sum_{r=1}^{R} & d\mathbf{a}_r \boxtimes (\mathbf{b}_r \boxtimes \mathbf{1}_M + \mathbf{1}_L \boxtimes \mathbf{b}'_r) \\ & + \mathbf{a}_r \boxtimes (d\mathbf{b}_r \boxtimes \mathbf{1}_M + \mathbf{1}_L \boxtimes d\mathbf{b}'_r) \end{aligned}$$

For the non multilinear model, we mainly answer two questions. Firstly, does the non bilinearity restore identifiability for the matrix case obtained by degenerating the first mode, *i.e.* with a single 2D measurement? If this were true, then measurements in fluorescence spectrometry could be done with a single experiment instead of repeating several experiments with different concentrations. Secondly, how does non-multilinearity affect exceptions to identifiability?

4.2. Matrix case

The next proposition answers the first question negatively. For simplicity, we have considered the matrix case $K = 1$ as if the factors $\mathbf{a}_r$ did not appear in the NLFD Jacobian, which is not true *stricto sensu*. The reason is that the arguments for rank deficiency of the non multilinear Jacobian are strictly identical when taking the first mode factors into account. In others words, one may set K = 1 instead of discarding first mode.

Proposition 3 *For the matrix case,* i.e. *with $\mathbf{a}_r$ playing no role in the NLFD, the non linear decomposition with rank $R > 1$ does not restore identifiability. Moreover, the rank of the Jacobian $\mathbf{J}_{NL}$ is upper bounded :*

$$rank(\mathbf{J}_{NL}) \leq R\,(L + M - 1)$$

with equality only when $R = 1$

Proof. In the matrix case, we do not have parameters $\mathbf{a}_r$ so that:

$$d\mathbf{e}(\boldsymbol{\theta}) = \mu \sum_{r=1}^{R} d\mathbf{b}_r \boxtimes \mathbf{1}_M + \mathbf{1}_L \boxtimes d\mathbf{b}'_r \stackrel{\text{def}}{=} \mu \mathbf{J}_{Exp} d\theta$$

Note that $\mathbf{J}_{Exp}$, of size $LM \times R(L+M)$, is the sum of identity matrices, and combinations of vectors of ones and zeros.

$$\mathbf{J}_{Exp} = [\, 0 \,|I_L \boxtimes \mathbf{1}_M + \mathbf{1}_L \boxtimes I_M| \ldots |I_L \boxtimes \mathbf{1}_M + \mathbf{1}_L \boxtimes I_M]$$

We can now deduce the Jacobian of the non linear model :

$$\begin{aligned} d\mathbf{x} &= \mathbf{J}_{NL}(\boldsymbol{\theta})\, d\boldsymbol{\theta} \\ \mathbf{J}_{NL}(\boldsymbol{\theta}) &= \mathbf{h}(\boldsymbol{\theta}) \boxdot (\mathbf{J}(\boldsymbol{\theta}) - \mu \mathbf{g}(\boldsymbol{\theta}) \boxdot \mathbf{J}_{Exp}(\boldsymbol{\theta})) \end{aligned} \tag{7}$$

where the Hadamard product between a vector and a matrix is the matrix defined as the iteration of the vector-vector product for every column of the matrix. Thus, the rank of the Jacobian $\mathbf{J}_{NL}$ is the rank of the right-hand side of the Hadamard product in (7). For better understanding of the structure of this matrix, we give the Jacobian for the case $L = M = 2$ and $R = 1$:

$$\mathbf{J} - \mu \mathbf{g} \boxdot \mathbf{J}_{Exp} = \begin{pmatrix} b_1 & 0 & c_1 - 2\mu b_1 c_1 & 0 \\ 0 & b_1 & c_2 - \mu b_1 c_2 & -\mu b_1 c_2 \\ b_2 & 0 & -\mu b_2 c_1 & c_1 - \mu b_2 c_1 \\ 0 & b_2 & 0 & c_2 - 2\mu b_2 c_2 \end{pmatrix}$$

Now one can notice that it is easy to cancel out, block by block, some Jacobian entries resulting from the exponential in (4) by using the L first columns. The involved coefficients should only be non-zero, which is generically true. With the notation of Proposition 1, with indices written modulo $(L+M)$,

$$\mathcal{C}_{M+i} \leftarrow \mathcal{C}_{M+i} + \mu c_i \mathcal{C}_i$$

With our example, it holds that $\mathbf{J} - \mu \mathbf{g} \boxdot \mathbf{J}_{Exp}$ has same rank as:

$$\begin{pmatrix} b_1 & 0 & c_1 - \mu b_1 c_1 & 0 \\ 0 & b_1 & c_2 - \mu b_1 c_2 & 0 \\ b_2 & 0 & 0 & c_1 - \mu b_2 c_1 \\ 0 & b_2 & 0 & c_2 - \mu b_2 c_2 \end{pmatrix}$$

We are then back to the multilinear case, since the matrix obtained this way has the same structure as the matrix in the CP decomposition case, and it is well known that the matrix congruent diagonalization problem admits an infinity of solutions (the SVD provides one solution among many others). Note that this is not true for $K > 1$, since the last cancellation is not valid anymore. This case is studied numerically in the next paragraph. □

4.3. NLFD exceptions to identifiability

It is well known that identifiability of CP decomposition has a few unexpected exceptions as presented in Section 3. The list is actually not exhaustive, but the other exceptions are either not relevant here because of the bound in Proposition 1, or deal with global uniqueness.

One of our contributions is to consider exceptions occurring in the frame of NLFD, and to show numerically that the defective unbalanced case is distorted but present, whereas the exception noted as defective earlier in Section 3 does not appear anymore. Computer results are reported in the tables beneath. Regardless of the number of iterations, the results were the same, so that we omit these data in the results. We insist that the results given here are generic, hence they are true in a dense subset of tensors, but may not apply to (rare) specific tensors, which have to be studied separately.

Because of space limitations, we present computer results in only a few cases to enlighten the previous discussion. For various tensors dimensions, we computed the difference between numerical rank and critical rank of the NLFD Jacobian. Simulations were run on tensors constructed with uniformly drawn factors, so that results presented here are of course meant generic, and very rare cases may not follow the rule. The following table gives the expected rank minus the computed rank of the Jacobian matrix for $R = \overline{R}$. Expected results are zeros, others are exceptions.

M	7			8			9			
$K\backslash L$	2	3	4	2	3	4	2	3	4	5
2	0	0	0	0	-5	0	-4	-5	-6	-7
3	0	0	0	-4	0	0	-4	0	0	0
4	0	0	0	0	0	0	-5	0	0	0
5	0	0	0	0	0	0	-6	0	0	0
6	0	0	0	0	0	0	0	0	0	0

If we were to characterize the exceptions, one would think the structure is very similar to the unbalanced exception for the CP decomposition, but with different boundaries. Moreover, no isolated exceptions where found for the NLFD in contrast to the defective case from section 3. From the complete data set, a lower bound $B(K, L, M)$ can thus be inferred, defining a boundary between balanced and unbalanced cases in the NFLD decomposition. The exact boundary was not determined, but we conjecture that for $M \geq (K-1)(L-1)+3$ or one of the two other symmetric inequalities verified, exceptions do not occur generically. This covers all practical cases in chemometrics.

5. CONCLUSION

We derived an easy approach to local identifiability. Properties from the literature were revisited, and extended to a non multilinear decomposition stemming from chemometrics. The general rule and its exceptions turn out to be different (compared to CP) for the non multilinear model under consideration.

REFERENCES

[1] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, pp. 455–500, 2009.

[2] P. Comon, "Tensors: a brief introduction," *IEEE Sig. Proc. Magazine*, vol. 31, no. 3, pp. 44–53, May 2014.

[3] X. Luciani, S. Mounier, R. Redon, and A. Bois, "A simple correction method of inner filter effects affecting FEEM and its application to the Parafac decomposition," *Chem. Intel. Lab. Sys.*, vol. 96, pp. 227–238, 2009.

[4] V. De Silva and L.-H. Lim, "Tensor rank and the ill-posedness of the best low-rank approximation," *SIAM J. Matrix Ana. Appl.*, vol. 30, pp. 1084–1127, 2008.

[5] C. Bocci, L. Chiantini, and G. Ottaviani "Refined methods for the identifiability of tensors" in *Annali di Matematica Pura ed Applicata*, pp. 1–12, 2013,

[6] J. B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions," *Lin. Alg. Appl.*, vol. 18, pp. 95–138, 1977.

[7] I. Domanov, "Study of canonical polyadic decomposition of higher-order tensors," Ph.D. dissertation, KUL, Leuven, Sep. 2013.

[8] I. Domanov, L. De Lathauwer, "On the Uniqueness of the Canonical Polyadic Decomposition of third-order tensors", SIAM Journal on Matrix Analysis and Applications, vol. 34, no. 3, pp. 855-903, 2013

[9] L. Chiantini and G. Ottaviani, "On generic identifiability of 3-tensors of small rank," *SIAM J. Matrix Ana. Appl.*, vol. 33, pp. 1018–1037, 2012.

[10] H. Abo, G. Ottaviani, and C. Peterson, "Induction for secant varieties of segre varieties," *Trans. AMS*, vol. 361, pp. 767–792, 2009.

[11] L. Chiantini and C. Ciliberto, "On the concept of k-secant order of a variety," *Journal of the London Mathematical Society*, vol. 73, pp. 436–454, 2006.

[12] P. Comon, J. M. ten Berge, L. De Lathauwer, and J. Castaing, "Generic and typical ranks of multi-way arrays," *Lin. Alg. Appl.*, vol. 430, pp. 2997–3007, 2009.

[13] X. Liu and N. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Sig. Proc.*, vol. 49, pp. 2074–2086, 2001.

[14] S. Sahnoun and P. Comon, "Tensor polyadic decomposition for antenna array processing," in *21st Int. Conf. Comput. Stat. (CompStat)*, Geneva, Aug. 19-22, 2014.

[15] P. Paatero, "A weighted non-negative least squares algorithm for three-way Parafac factor analysis," *Chem. Intel. Lab. Sys.*, vol. 38, pp. 223–242, 1997.